Environmental and Molecular Mutagenesis 32:188–191 (1998)

Transition Mutations at CpG Dinucleotides Are the Most Frequent In Vivo Spontaneous Single-Base Substitution Mutation in the Human *HPRT* Gene

J. Patrick O'Neill[1]* and Barry A. Finette[2]

[1] *Department of Medicine, University of Vermont College of Medicine, Burlington, Vermont*
[2] *Department of Pediatrics, University of Vermont College of Medicine, Burlington, Vermont*

Key words: CpG dinucleotides; *HPRT* mutation; human cells; Lesch-Nyhan syndrome; newborns

Studies of human genetic disorders have suggested that 5-methylcytosine moieties in DNA may act as endogenous mutagens due to the increased deamination of C when it is methylated in CpG dinucleotide sequences. This deamination of C to yield T can result in a DNA replication mediated C/G → T/A transition mutation. Fully 32% of human germline single-base substitution (SBS) mutations in the LDL receptor gene and 50% of germline SBS mutations in the *p*53 gene (Li-Fraumeni syndrome) are found in CpG dinucleotide sequences [Rideout et al., 1990; Cooper and Krawczak, 1990]. Deamination of C yields U, while deamination of 5-methylcytosine yields T. These C to T (or G to A) transitions have been proposed to be due to the decreased rate of repair of T:G mismatches versus U:G mismatches [Schmutte et al., 1995]. It is this difference in repair which causes 5-methylcytosine to stand out as a ''hot spot.''

We recently determined the *HPRT* mutations in two unrelated families each with a Lesch-Nyhan syndrome-affected male and found that both contain a C_{151} to T transition mutation by cDNA sequencing (Table I). In the first family ($HPRT_{Estrie}$), we determined that both the mother and the sister of the affected boy showed the elevated thioguanine-resistant (TG^r) mutant frequency expected for carrier females [Skopek et al., 1990; Hunter et al., 1996] and both showed the C_{151}GA to TGA mutation in the mutant allele cDNA. In the second family ($HPRT_{Sagamie}$), the proband and his heterozygous mother showed the same mutation. This mutation occurs at a CpG dinucleotide and results in a change from an arg to a nonsense (chain-terminating TGA) codon. Previous studies have suggested that this site might be a hot spot for germline *HPRT* mutation [Sege-Peterson et al., 1993; Alford et al., 1995; Fujimori et al., 1995]. This observation of two new families with the $C_{151} \rightarrow T$ mutation led us to investigate the distribution of CpG dinucleotides within the *HPRT* gene and the frequency of mutations at these sites.

Germ cell mutations in the X-chromosome *HPRT* gene clinically result in either Lesch-Nyhan syndrome (severe loss of enzyme activity) or gout (partial loss of enzyme activity). The human *HPRT* gene has a 1.4 kilobase mRNA containing a 657 base-coding region. The dinucleotide sequence CpG would be expected to appear 27 times by random distribution, but is only found eight times; i.e., it is reduced by 70%. It has been proposed that this is the result of mutational pressure against this highly mutable dinucleotide [Kricker et al., 1992]. The eight CpG dinucleotides are shown in Table II, where the A of the AUG initiation codon is designated base 1. These eight CpG dinucleotides present 16 possible SBS (transition) mutations: CpG → TpG or CpG → CpA. Four of these result in no predicted amino acid change, ten in a missense mutation, and two in a chain-terminating (TGA) codon. Five of the possible 12 mutations have been found to occur in in vivo or in vitro studies and these are shaded in Table II. Table III summarizes the distribution of SBS (transition) mutations at these five sites, as reported in the Cariello *HPRT* database [Cariello, 1997]. Two of these mutations ($CpG_{481} \rightarrow CpA$ and $CpG_{509} \rightarrow CpA$) have been reported only in human cells in in vitro studies, the former

Contract grant sponsor: the National Institute of Child Health and Human Development; Contract grant number: 1K11HD01010; Contract grant sponsor: the Charles H. Hood Foundation; Contract grant sponsor: NIH; Contract grant number: 5R01CA68288. Support does not imply agreement with the views expressed in this commentary.

*Correspondence to: J. Patrick O'Neill, University of Vermont Genetics Laboratory, 32 North Prospect Street, Burlington, VT 05401. E-mail: patrick.oneill@uvm.edu

Received 10 January 1998; Accepted 13 April 1998

TABLE I. Analysis of Potential *HPRT* Mutation Carrier Females

Sample	Individual	*HPRT* Mutant frequency	Mutation
(P52) Estrie	Proband	(unavailable)	
	Mother	3.7×10^{-2}	$C_{151}GA \rightarrow TGA$
	Sister	1.6×10^{-2}	$C_{151}GA \rightarrow TGA$
(P55) Sagamie	Proband	0.98	$C_{151}GA \rightarrow TGA$
	Mother	3.5×10^{-2}	$C_{151}GA \rightarrow TGA$

TABLE II. CpG Dinucleotide Sequences in the Human *HPRT* Gene

cDNA sequence	Mutation	Predicted amino acid change	Mutation	Predicted amino acid change	C/NC*
GC_5G	$C_5 \rightarrow T$	ala → val	$G_6 \rightarrow A$	(none)	NC
$C_{10}GC$	$C_{10} \rightarrow T$	arg → cys	$G_{11} \rightarrow A$	arg → leu	NC
$GGC_{21}GTC$	$C_{21} \rightarrow T$	(none)	$G_{22} \rightarrow A$	val → ile	NC-NC
$GTC_{24}GTG$	$C_{24} \rightarrow T$	(none)	$G_{25} \rightarrow A$	val → met	NC-C
$C_{142}GT$	$C_{142} \rightarrow T$	arg → cys	$G_{143} \rightarrow A$	arg → his	NC
$C_{151}GA$	$C_{151} \rightarrow T$	arg → Stop	$G_{152} \rightarrow A$	arg → gln	NC
$GTC_{480}GCA$	$C_{480} \rightarrow T$	(none)	$G_{481} \rightarrow A$	ala → thr	NC-C
$C_{508}GA$	$C_{508} \rightarrow T$	arg → Stop	$G_{509} \rightarrow A$	arg → gln	NC

*Conserved (C) or not conserved (NC) amino acid [Lambert et al. 1992].

in T-lymphocytes treated with nitrosopyrene, the latter in GM130 lymphoblasts treated with ethylnitrosourea and in fibroblasts treated with ultraviolet light. A third ($CpG_{143} \rightarrow CpA$) has been reported four times, twice in T-lymphocytes (one in an untreated culture in an in vitro study and one as an in vivo-arising mutation in an unexposed, nonsmoking male) and twice in humans with inherited *HPRT* partial deficiency (resulting in gout). The other two mutations ($CpG_{151} \rightarrow TpG$ and $CpG_{508} \rightarrow TpG$) have been reported frequently in both in vitro and in vivo studies and represent 3.1% (43/1,403) of the total SBS mutations reported in the total database [Cariello, 1997]. All five sites represent 3.6% of the total SBS mutations (Table III, Line A).

Transition mutations at CpG dinucleotides account for 15.7% of the Lesch-Nyhan (CpG_{151} and CpG_{508}) and 7.7% of the gout (CpG_{143}) mutations in the *HPRT* database (Table III; Line D). Similar results of 14.0% and 16.7% are found in two other compilations of Lesch-Nyhan mutations which partially overlap the *HPRT* database entries (Table III; Lines E and F). Taken together, these CpG mutations are the most frequent germline *HPRT* mutations. These results are consistent with previous reports of the high frequency of CpG mutations in other inherited diseases [Rideout et al., 1990].

In vivo-arising somatic cell SBS mutations in T-lymphocytes comprise 54.1% (253/468) of the total *HPRT* mutations and occur at a frequency of 5.1% (13/253) at the three CpG sites also seen in germ cell mutations (Table III; Line C). Again, this is the most frequent somatic *HPRT* mutation in adults. T-lymphocyte *HPRT* mutation frequencies have been shown to increase with age [Robinson et al., 1994] and the role of cell proliferation in mutation rate calculations has been discussed [Green et al., 1995]. The rate of increase in mutant frequency is greater in children than in adults, consistent with the higher level of T-cell proliferation in children [Finette et al., 1994]. The frequency of mutations in newborns (cord blood samples) probably presents the greatest contribution of active cell proliferation to mutation induction which results in an elevated mutation rate in newborns as compared to adults [Green et al., 1995]. One prediction that results from the hypothesis that mutations at CpG dinucleotides are DNA replication-dependent is that SBS mutations in newborns should occur predominantly at these sites as true spontaneous or endogenous mutation events. Table III, Line G presents preliminary data that this prediction is true. Although SBS mutations comprise only 23% (18/77) of the total *HPRT* mutations in newborns (with the majority being large structural alterations), six of these 18 SBS mutations (33.3%) are transition mutations at two CpG dinucleotide sequences ($C_{151}pG$ and $C_{508}pG$).

These observations suggest that deamination of 5-methylcytosine to T at CpG dinucleotide sequences and subsequent replication resulting in C/G → T/A transition mutations are the most frequent mutation event in dividing cells in vivo, both somatic and germ cells. The decline

TABLE III. *HPRT* Mutations at CpG Dinucleotides

Dataset	Total number of mutations	Number of single-base substitutions	Number of mutations at					Total (% of SBS)
			C_{151}	C_{508}	G_{143}	G_{481}	G_{509}	
A. Total[a]	2224	1403	20	23	4	1	2	50 (3.6%)
B. In vitro[a]								
Total	1555	1024	10	11	1	1	2	25 (2.4%)
''Induced''	1366	914	7	10	0	1	2	20 (2.2%)
''Spontaneous''	189	110	3	1	1	0	0	5 (4.5%)
C. In vivo								
Somatic[a] (Spontaneous and smoker)	468	253	4	8	1	0	0	13 (5.1%)
D. In vivo								
Somatic (newborn)	77	18	3	3	0	0	0	6 (33.3%)
E. Germinal[a]								
LN	102	70	7	4	0	0	0	11 (15.7%)
Gout	33	26	0	0	2	0	0	2 (7.7%)
F. Germinal[b]								
LN	65	43	3	3	0	0	0	6 (14.0%)
Partial	20	18	0	0	0	0	0	
G. Germinal[c]	41	24	2	2	0	0	0	4 (16.7%)

[a]Cariello, 1997.
[b]Scully et al., 1992.
[c]Alford et al., 1995.

in in vivo somatic cell CpG mutation frequency in adults compared to newborns (5.1% vs. 33.3%, respectively) probably reflects the decline in cell proliferation in the T-lymphocyte population after birth. However, since the age-related increase in mutant frequency from newborn to adult is approximately 15-fold compared to the 6-fold decline in the frequency of mutation at these CpG sites in adults, these specific mutations appear to continue to occur after birth, although at a lower frequency. These mutations may represent true spontaneous mutations due to the endogenous mutagenic potential of 5-methylcytosine deamination, coupled with DNA replication of the resulting T:G mismatch, to yield C/G → T/A transition mutations. The frequency of these mutations in newborns may provide a measure of the inherent spontaneous mutation events and thus allow determination of induced events beyond this background.

ACKNOWLEDGMENTS

The two Lesch-Nyhan families were studied in collaboration with Drs. S.B. Melancon and L. Dallaire of the Universite' de Montreal, Montreal, Quebec, Canada.

REFERENCES

Alford RL, Redman JB, O'Brien WE, Caskey CT (1995): Lesch-Nyhan syndrome: Carrier and prenatal diagnosis. Prenat Diagn 15:329–338.

Cariello NF (1997): ''Release 6 of the *HPRT* Database.'' Mutabase Software, Durham, NC.

Cooper DN, Krawczak M (1990): The mutational spectrum of single base pair substitutions causing human genetic disease: Patterns and predictions. Hum Genet 85:55–74.

Finette BA, Sullivan LM, O'Neill, JP, Nicklas JA, Vacek PM, Albertini RJ (1994): Determination of *HPRT* mutant frequencies in T-lymphocytes from a healthy pediatric population: Statistical comparison between newborn, children and adult mutant frequencies, cloning efficiency and age. Mutat Res 308:223–231.

Fujimori S, Tagaya T, Yamaoka N, Saito H, Kamatani N, Akaoka I (1995): Direct evidence for a hot spot of germ line mutation at *HPRT* locus. In: Sakota A, Taylor A (eds): ''Purine and Pyrimidine Metabolism in Man VIII.'' New York: Plenum Press, pp 679–682.

Green MHL, O'Neill JP, Cole J (1995): Suggestions concerning the relationship between mutant frequency and mutation rate at the *HPRT* locus in human peripheral T-lymphocytes. Mutat Res 334:323–339.

Hunter TC, Melancon SB, Dallaire L, Taft S, Skopek TR, Albertini RJ, O'Neill JP (1996): Germinal *HPRT* splice donor site mutation results in multiple RNA splicing products in T-lymphocyte cultures. Somat Cell Mol Genet 22(2):145–150.

Kricker MC, Drake DW, Radman M (1992): Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. Proc Natl Acad Sci USA 89:1075–1079.

Lambert B, Marcus S, Anderson B, Hou S-M, Steen AM, Hellgren D (1992): Missense mutations and evolutionary conserved amino acids at the human *HPRT* locus. Pharmacogenetics 2:329–336.

Rideout WM, Coetzee GA, Oleimi AF, Jones PA (1990): 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and *p*53 gene. Science 249:1288–1290.

Robinson DR, Goodall K, Albertini RJ, O'Neill JP, Finette B, Sala-Trepat M, Moustacchi E, Tates AD, Beare DM, Green MHL,

Cole J (1994): An analysis of *in vivo HPRT* mutant frequency in circulating T-lymphocytes in the normal human population: A comparison of four datasets. Mutat Res 313:227–247.

Schmutte C, Yang AS, Beart RW, Jones PA (1995): Base excision repair of U:G mismatches at a mutational hotspot in the *p53* gene is more efficient than base excision repair of T:G mismatches in extracts of human colon tumors. Cancer Res 55:3742–3746.

Sculley DG, Dawson PA, Emmerson BT, Gordon RB (1992): A review of the molecular basis of hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) deficiency. Hum Genet 90:195–207.

Sege-Peterson K, Chambers J, Page T, Jones OW, Nyhan WL (1992): Characterization of mutations in phenotypic variants of hypoxanthine phosphoribosyltransferase deficiency. Hum Mol Genet 1(6):427–432.

Skopek TR, Recio L, Simpson D, Dallaire L, Melancon SB, Ogier H, O'Neill JP, Falta MT, Nicklas JA, Albertini RJ (1990): Molecular analyses of a Lesch-Nyhan syndrome mutation ($HPRT_{Montreal}$) by use of T-lymphocyte cultures. Hum Genet 85:111–116.

Accepted by—
T. Skopek